A novel vertical handover mechanism for media streaming
in heterogeneous wireless architectures

LÁSZLÓ BOKOR, LÁSZLÓ LOIS, CSABA A. SZABÓ, SÁNDOR SZABÓ

Budapest University of Technology and Economics, Department of Telecommunications
{bokorl, lois, szabo, szabos}@hit.bme.hu

Keywords: heterogeneous wireless environment, streaming media delivery, WiFi, 2.5G and 3G mobile

The paper presents a novel vertical handover mechanism which aims at assuring streaming media services in a heterogeneous network environment where the subscribers are roaming among different wired/wireless access systems including ADSL, WiFi, 2.5G and 3G cellular and WiMAX. The handover scheme provides seamless connectivity during roaming, with adapting the quality of the delivered media stream to the changes of the network characteristics and to the capabilities of a wide variety of devices. The paper presents general design considerations, focuses on introducing the operational behaviour of the novel vertical handover method, its implementation and evaluation in a heterogeneous access network testbed, as well as discusses some aspects of gaining higher level features based on our proposal.*

1. Introduction

The key issue of any handover management mechanism designed for heterogeneous architectures is the efficient management of all kind of transitions between access networks, called briefly mobility (we are using this term in a broader sense, denoting movement from one access network to another, regardless of the velocity of the user, thus including real or nomadic mobility, etc.). Traditional mobility management is hedged in providing *terminal mobility*. This kind of mobility allows a mobile node to maintain ongoing communication or commence/receive incoming session requests independently from its network point of attachment. (Note that there is a subset of terminal mobility which ensures handling only new sessions after changes of networks. Providing this subset requires only dynamic DNS and DHCP functions.) The development, deployment and convergence of different wired and wireless technologies introduced several new mobility types which can be grouped into two main categories. On one hand, there is a device-centric, low-level mobility including *ad hoc mobility* (mobile devices are routable in ad hoc networks) and *mode mobility* (devices can switch between ad hoc and infrastructure modes). On the other hand, we can talk about user-centric, high-level mobility consisting of *personal mobility* (users are globally reachable at different scenarios: one address for many different devices or many addresses reaching one device), *session mobility* (active sessions are switchable between terminals) and *service mobility* (personalized services can be maintained while moving and/or changing devices/ISPs).

Supporting mobility between different types of access networks (e.g. UMTS to WLAN) is called "vertical handover" in order to distinguish it from the usual "horizontal handover" (the migration of mobile nodes between homogeneous networks, e.g. UMTS to UMTS) often occurring in a mobile operator's network whenever a user leaves the radio cell of a base station and enters a neighboring cell [15]. (Note that the difference between the terminology of horizontal and vertical handover is vague. For example, a handover from an 802.11b WLAN AP to an 802.11g AP link may be considered as either a vertical or a horizontal handover, depending on the point of view.)

In a media streaming delivery architecture, built on a heterogeneous architecture, managing vertical handovers during mobility is necessary for three main reasons. The first is obvious: while moving, the user approaches the boundary of the coverage area of the actual network or part of the network (e.g. a radio cell) so that the bit error rate, packet loss or any other QoS parameter becomes too high and an *automatic* (unplanned) vertical handover is to be performed. The second reason is when the user wants to change the current manner of network attachment intentionally because there is a better way available to support better QoS parameters for the media (consequently, often for higher price). This latter case is called *user-initiated* vertical handover. Third, when a load balancing mechanism of an overlay network is able to distribute the overall network load in order to optimize the performance of each individual network. This is called *network management-initiated* vertical handover which allows more efficient resource utilization for service providers.

The main novelty of our handover management approach lies in transparently and effectively supporting the close incorporation and tight integration of advanced

** Based on "Testbed of a Novel Media Streaming Architecture for Heterogeneous Wireless Environment", by László Bokor, László Lois, Csaba A. Szabó, Sándor Szabó, published in the Proceedings of the 3rd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TridentCom2007), Orlando, Florida, USA.*

vertical mobility management, multimedia processing and transmission, quality of service, adaptation to different network and device capabilities, digital rights management and flexible quality based billing. Using our vertical handover mechanism, all of these features and enhancements can easily be merged into a single integrated and transparent architecture by granting the capability of moving across a wide range of loosely-coupled access networks.

The paper presents an overview of the design, implementation and evaluation of this novel handover mechanism for media streaming in heterogeneous wireless architectures providing the features introduced above. The paper is organized as follows. Our main design choices are described in Section 2. In Section 3, the mechanisms to implement and test vertical handover management across heterogeneous access networks are described and test results are presented and evaluated. Section 4 deals with the integration possibilities of different features into a single media streaming architecture based on our vertical handover scheme. The paper concludes with a summary in Section 5.

2. Conceptual overview

The main requirements towards a mobility handling mechanism supporting vertical handovers for media streaming in heterogeneous architectures are as follows:

a) Efficient location management.

The mobile device should always be accessible by a static identifier regardless of its current location. The following challenges should be overcome:

- Reduction of signaling overhead and latency
- Guaranteed QoS in different access networks
- Intelligent decision algorithms controlling the MN's behavior in overlapping areas of heterogeneous wireless networks

b) Handling wide variety of mobility.

The mobility management solution should allow as many mobility types (terminal, personal, etc.) as possible.

c) Transparent and seamless handovers.

The change between different networks should not cause considerable data loss, the transition itself should not last long and the long-term connection-oriented protocols should run in a seamless way. The following challenges are to be solved:

- Reduction of signaling overhead and latency
- Guaranteed QoS during the handover procedures
- Scalability, efficient utilization of resources, reliability, robustness

d) Infrastructure-less solution.

The mobility management solution should be located at the boundaries of the network so that no or only minor changes are required in the service providers' networks.

We have evaluated all of these requirements in [2], in order to choose an appropriate handover management approach fitting into the proposed streaming media architecture. We have assessed our criteria and compared the parameters/attributes of mobility management methods at different layers of the TCP/IP architecture. We have concluded, that network and application layer mobility (Mobile IPv6 [13] and Session Initiation Protocol [6]) are the most promising approaches for our purposes. Considering these two methods we have pointed out that in case of terminal mobility Mobile IPv6 is the most general and effective solution but requires more complex infrastructure than SIP-based mobility.

Personal mobility cannot be achieved using only network layer methods; however SIP forking proxies in the application layer easily bypass this issue. Session mobility is not supported by Mobile IPv6 (however IPv6 anycasting could have the potential to provide this feature [3]) in contrast to SIP, which allows session mobility by explicitly transferring a session to another destination. Service mobility is achievable based on Mobile IPv6 (by using subsidiary components to keep service definitions updated) but SIP offers built-in mechanism for synchronizing service definitions and other configuration elements.

Besides that SIP is able to provide terminal, personal, session and service mobility, it supports the widest range of applications from VoIP, Internet conferencing and presence to instant messaging and event notification. SIP was originally proposed by IETF to establish, modify and release sessions in all-IP networks and then gained support by 3GPP to perform signaling tasks in IP Multimedia Subsystem (IMS) [7,8].

Now SIP is the basic protocol of the 3GPP IMS, and IMS is also incorporated in ETSI's NGN architecture. The SIP-based IMS defines an overlay architecture on the top of any 3G packet switched core network comprising the key technologies of the future's converged, service and application oriented networks [9,10].

With regard to the current trends in telecommunications, it is obvious that all players of the future mobile communication market need an easy-to-use instrument for quick integration and examination of new services and applications even from third parties. IMS seems to be that generic instrument and that was our most important reason to choose SIP and the application layer-located approach as the basis of our handover management proposal, despite the fact that an integrated and/or hybrid MIPv6-SIP architecture could combine the benefits of each protocol [4,5].

3. The proposed handover management mechanism

A. Components and basic operation

To provide a generic solution, the proposed mechanism separates media functions and mobility functions by segregating the following operational components *(Figure 1):*

Media Player, SIP Streaming Client, Media Server, SIP Streaming Server, SIP Server.

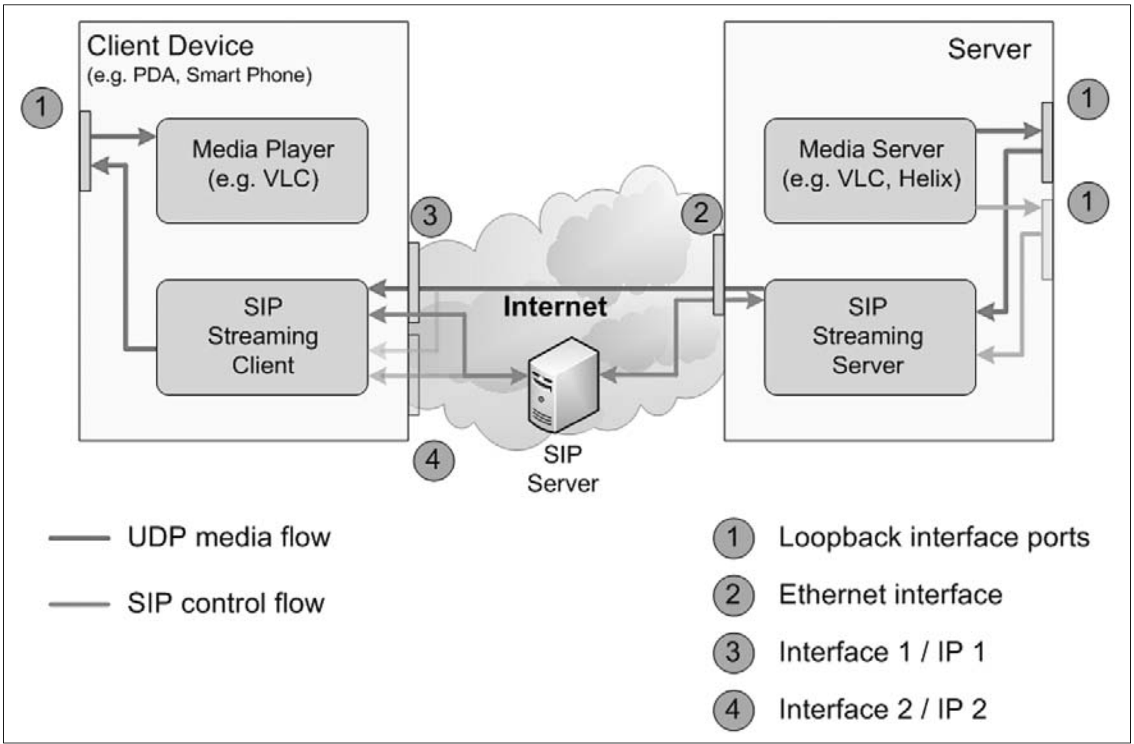


Figure 1.

SIP-based streaming architecture for heterogeneous environment

Any type of media server (e.g. HELIX, VLC) and player (e.g. VLC) can be used, which is capable of handling UDP/RTP network streams. The only task of the media server-player modules is to create and play back the media content. Any other function is completely independent from them, making their functions transparent and the implementation portable and open. The separate module-pair of SIP Streaming Client and Server is responsible for transporting the UDP/RTP media stream across various networks, by utilizing a SIP based user authentication and mobility management protocol. The SIP Server handles the standard functions of controlling the SIP-based communication.

The basic idea is the following: the mobile terminal (Client Device) registers its current IP address on an access network, and after successful authentication and registration, the SIP Streaming Server forwards the stream of the Media Server to the mobile terminal's registered IP address. The SIP Streaming Client captures the UDP/RTP media stream, and forwards it to a local UDP port on the mobile terminal. The Media Player – which should be able to accept streams from a defined UDP port – connects to this local port, and plays back the content. Upon handover, the mobile terminal registers its new IP address (using the functions of SIP Streaming Client) to the SIP Streaming Server, and the server transfers the media stream to the new address. The media player running on the client device is not aware of the handover event.

Notice the fact of separating the UDP/RTP based media and SIP signaling messages, according to NGN and IMS concepts [11,12]. We have to highlight that one of the most important advantages of utilizing SIP protocol in the proposed manner is the ease of integration with IMS system, as an AS (Application Server).

B. SIP streaming client and server

The main component of the described handover management method is the SIP Streaming Client. This module integrates the following functions *(Figure 2)*:

- *Forwarder:* Transmits the media stream from a given interface to the loopback port where the Media Player can reach and play back.
- *SIP Communicator:* Handles the SIP communication by managing SIP requests and responses.
- *Interface:* Monitors active interfaces/networks (listed in the user's preference file given by the GUI), and measures the QoS of present connections based on packet loss and round-trip time. Every interface owns a state machine describing its actual status (see Figure 7).
- *Connection Manager:* Selects the active connection based on the QoS measurements made on the interfaces, and upon a given threshold, initiates the handover and all concerning SIP procedures automatically. In addition to the architectural support, an efficient decision algorithm is also implemented for fast selection of connection to perform seamless, or nearly seamless vertical handover. (Note that in our testbed the threshold values can be predefined or dynamically calculated by the decision algorithm.)
- *Utility:* A little application which allows NAT traversal.
- *Graphical User Interface:* The Graphical User Interface allows the users to define the precedence of the potential network interfaces, presents information about the current connection and active interfaces, and helps to execute user-initiated handovers.

The main task of the decision algorithm operating in the SIP Streaming Client is to pick out the most appropriate connection (interface) from the user-designated

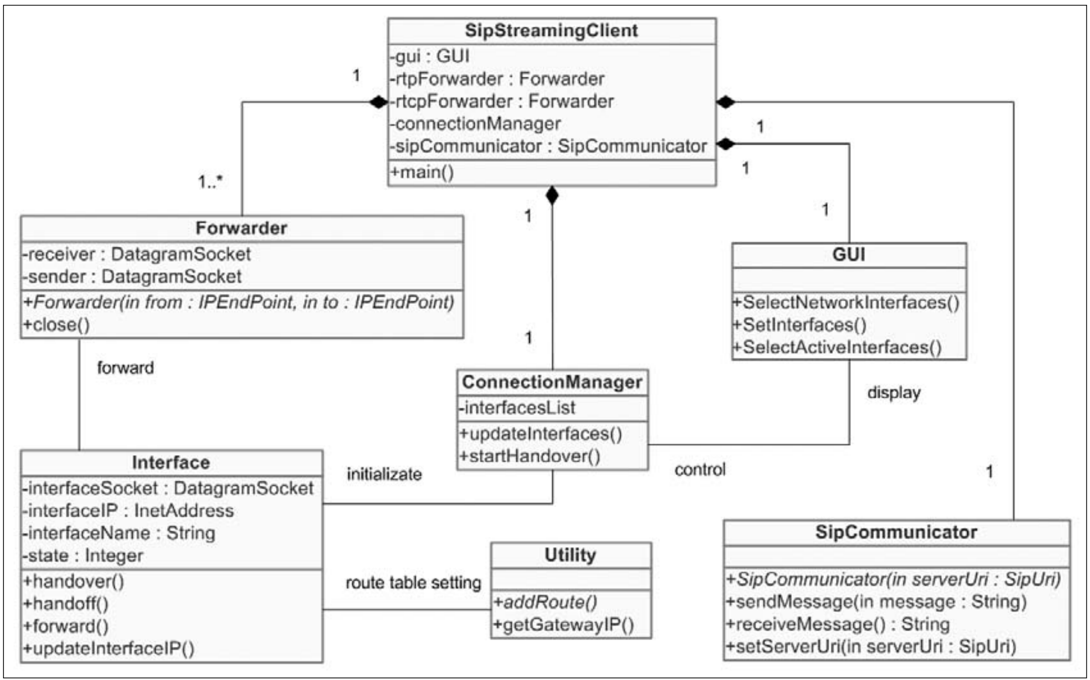


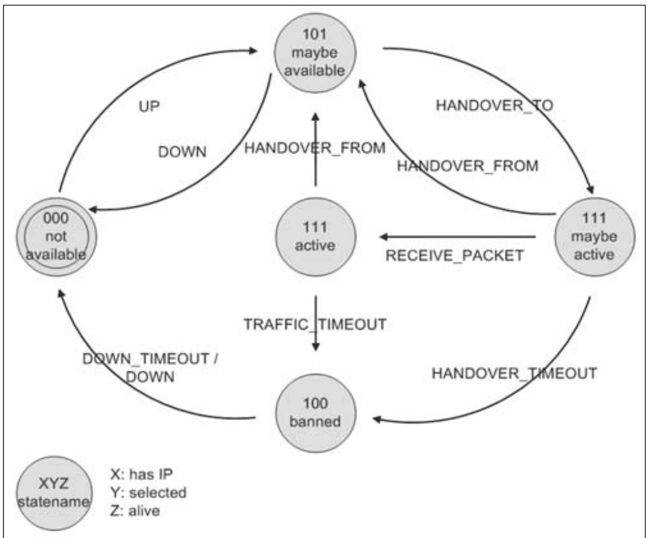
Figure 2.

UML class diagram of SIP Streaming Client

ones. In order to achieve this function, the decision engine continuously gets and evaluates jitter and packet loss measurement results (referenced among the SIP Streaming Server and the active interfaces), and continuously calculates and updates the threshold value of beginning a transition. At this phase no physical layer information is used as an input parameter for the algorithm: the decision is based only on IP level attributes, calculated/predefined threshold values, different states of interfaces and on user-preferences. This enables simpler implementation and build-up while keeping the effective operation and adequate performance as well.

Because the need of a handover transition doesn't depend only on calculated or predefined threshold values but on different states of interfaces and on user-preferences as well, a state machine was defined to maintain the altering conditions of every interface and to control the process *(Figure 3)*.

Figure 3. State machine of an interface



The SIP Streaming Server is responsible for managing the multimedia sessions by relaying between a media player and server and for transmitting the media streams always to the actual addresses of the clients by parsing incoming SIP signaling messages and forwarding the media content from the local port towards the mobile node's actual location.

Compared to the SIP Streaming Client, this software module does not contain any special intelligence: mainly it is a simple SIP interpreter driven by SIP commands in order to achieve a dynamic, transparent and media server independent source of media content.

C. Experimental results

To evaluate our vertical handover method designed for media streaming in heterogeneous environment, we have set up an experimental testbed consisting of a loosely-coupled UMTS-WLAN-LAN architecture *(Figure 4)*. This experimental testbed combines several independent IP-based networks, including T-Mobile Hungary's UMTS architecture, IEEE 802.11a/b/g WLAN network, and BME-HT's* LAN. All of these technologies are integrated by a common, IP-based operation, however below IP each access network has its own protocol stack, differing characteristics and attributes.

The 3G/HSDPA cellular UMTS network and all related infrastructure are organic part of T-Mobile Hungary's production network. This component enables services at maximum 1.5 Mbps download and 384 Kbps upload speeds with RTTs around 80 ms. The WLAN connectivity is a separate sub-network of our testbed. There are IEEE 802.11a/b/g compatible Linksys WAP55AG Access Points interconnected. This exclusive, local wireless access enables Mobile Nodes to transmit/receive data up to 54 Mbps with RTTs around 10 ms.

** Budapest University of Technology and Economics, Department of Telecommunications*

The LAN component is an IEEE 802.3 compatible Ethernet connection established at BME-HT. This connectivity is used by cable link enabling high-speed wired networking up to 100 Mbps with negligible RTTs (around 1 ms).

For accessing this heterogeneous network setup, Mobile Nodes have been equipped with separate interfaces for every access component, and with a Java implementation of our SIP Streaming Client software. A MN's hardware is based on Fujitsu-Siemens Lifebook C1320D (1.73 GHz Pentium M processor, 256 MB RAM) with BroadCom NetXtreme Gigabit Ethernet interface, Edimax ZD1201 USB WLAN adapter, and Globetrotter 3G+ PCCard Modem for UMTS connectivity.

The testbed's Media Server and SIP Server is a standard PC with a 3 GHz Pentium 4 processor and 512 MB RAM, running a Java implementation of our SIP Streaming Server and SIP Proxy/Registrar.

Based on this heterogeneous testbed, experimental surveys can be performed in order to observe Mobile Nodes while they perform seamless roaming among different access technologies, and maintain ongoing media connections. The first experiments were focused on vertical handover management, the results obtained in these scenarios using the basic testbed topology are presented in *Figures 5, 6, 7 and 8*.

The amount of time required by the handover among different network types, and the resulting packet loss were measured to present the performance characteristics of our vertical handover-enabled media streaming architecture. We gathered measurement data using preset values based and decision algorithm aided connection detection, and performed several trials in the same operating conditions for three different classes of media traffic (48, 192, 385 Kbps) in three different upward vertical handover scenarios (LAN to WLAN, LAN to UMTS, WLAN to UMTS). The values in the tables represent average values obtained on the mobile node (moving client of a streaming application) across five repetitions of each test. The handovers were initiated by simulated total link failure (e.g. disconnection of the physical medium); in more realistic situations even better results are expected.

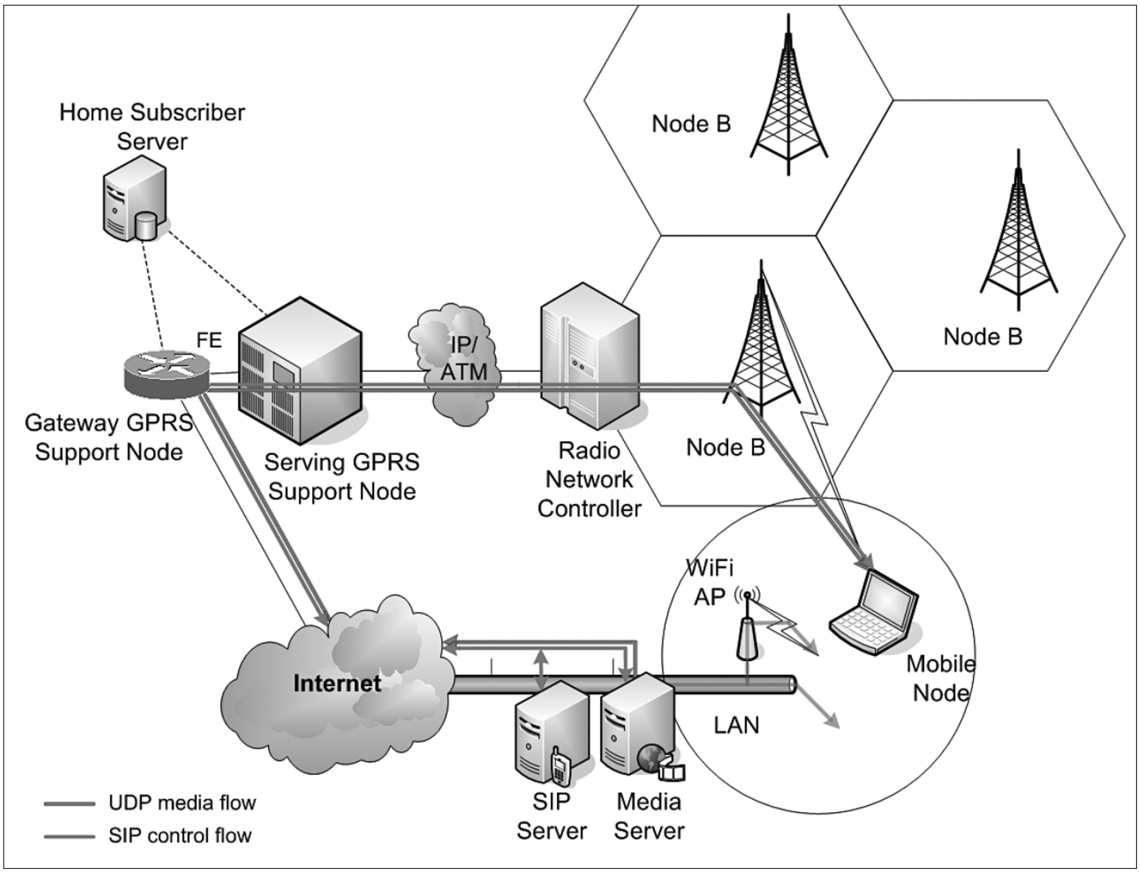
In case of user initiated handover the system provides no packet loss during the transition (soft-handoff), thus no visible deterioration of the media stream can be observed unlike in case of automatic (unplanned) handovers where the packet loss of an ongoing media stream depends on three main factors:

– Bit rate of ongoing media traffic;
– Timeout of detecting the network failure on the current link;
– RTT of the new network (i.e. delay of SIP signaling messages managing the handover procedures).

The comparison of packet loss measurement results presented in *Figure 5 and 6* shows that the decision algorithm aided vertical handover outperforms the preset values based method by dynamically and effectively adapting the SIP Streaming Client to the continuously varying network conditions. Results of vertical handover latency measurements in *Figure 7 and 8* confirm the

Figure 4.

Basic topology for the measurements



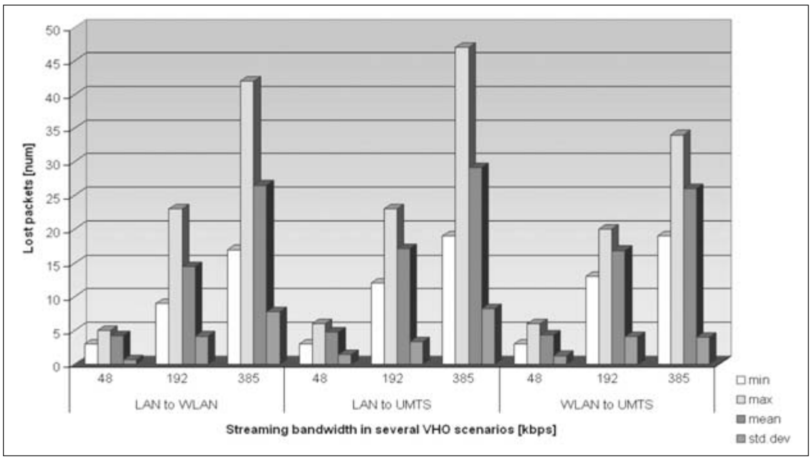


Figure 5.
Number of packets lost
(Preset values based connection detection)

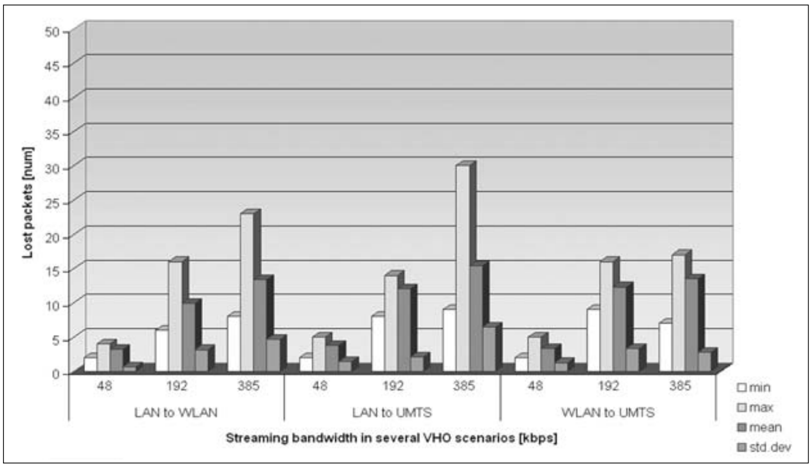


Figure 6.
Number of packets lost
(Decision algorithm aided connection detection)

above statement and show that the VHO latency will be less than 2.5 s in all the examined scenarios even without using adaptive clients based on decision algorithm aided connection detection.

4. Higher level features

In a multi-platform access network environment, the user has several physical connections to access the Internet, hence the same resource could be even accessed simultaneously via different wireless networks. Based on a transparent and effective vertical handover mechanism, this opportunity could be utilized to achieve higher quality service, i.e. faster download or higher quality media streaming solution by using all access networks simultaneously or selecting the best access network(s) dynamically. The main components of such an integrated media streaming architecture for heterogeneous environment are the following:

- Media delivery subsystem
- Network access and handover module
- Bit-rate and resolution selection and ranking module
- Security and accounting module

However, all of these modules should cooperate with each-other to achieve a higher level service. The cooperation of the main modules of such a media streaming system based on our transparent vertical handover mechanism is shown in *Figure 9 (on the next page)*.

The media delivery subsystem provides important information such as:

- decoded video and audio quality and buffer fullness,
- available bit-rates and resolutions of the current media content,
- aviable bit-rate switching positions.

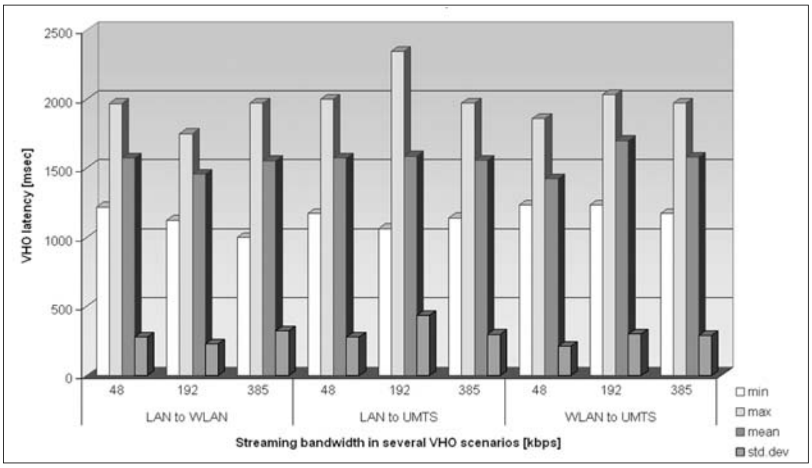


Figure 7.
VHO latency
(Preset values based connection detection)

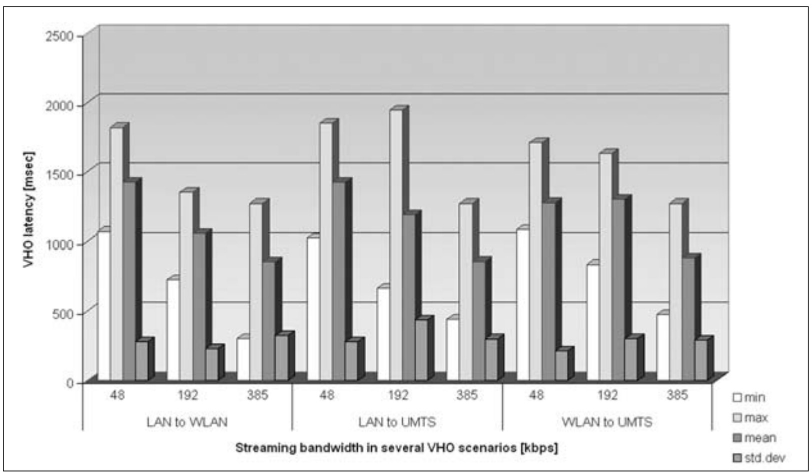


Figure 8.
VHO latency
(Decision algorithm aided connection detection)

While the media server knows when the bit-rate switching could be performed without any transient effects in the reconstructed video, the network access module performs the switching from one access network to another in case of need. By communicating the new destination IP address, the new bit-rate and resolution to the media server and client, the media presentation becomes continuous (in exception of an abrupt network loss) when the client sends retransmission requests and uses enough long look-ahead buffer at the input of the media decoder. Since the network access units have different IP addresses, the network access and handover client forwards the media stream arriving on the active port to a fixed virtual port referring to the media client.

The different client capabilities are the input of the bit-rate selection and ranking module, and the other input of this module is the current state of each access network. The client capabilities can be evaluated at start-up, but the state of the access network is reported regularly by the network access and handover module. The selection of the new access network and the new bit-rate is done by the bit-rate selection and ranking module, but the bit-rate could also be changed while the access network remains the same, i.e. in case of remarkable alteration of the packet loss statistics.

In a system based on our vertical handover management scheme, the decision of the stream switching can be based on the client's measurements, since the available network connections are handled by the SIP client and the server has little to do with the possible connections of every client. Based on the measured parameters (current packet loss rate and the access network type), the optimal bandwidth can be estimated and the ranking of the access networks are made. Based on the ranking, the optimal bandwidth of the best connection and the decoder properties (i.e. screen size, decoding speed), the best bandwidth/quality version of the content is determined and the switching is carried out in case of need.

The ranking of the connections can be based on several measurements or pre-defined values, such as:

– the availability of the connection, even by measuring the field strength at the receiver's front-end if possible, or upper layer parameters (e.g. packet loss rate),
– the expected available/achievable bandwidth of the connection,
– the expected or actual packet loss rate for the idle or active connections, respectively,
– the expected video and audio quality of the media streaming over the connection,
– cost of the connection.

The ranking and the selection of the best connection is re-evaluated upon:

– degradation of parameters of the active connection,
– improvement of parameters of an idle connection.

With the proper ranking method, the media streaming system can operate on the active connection close to the optimal single-connection configuration on that connection. Since our proposal handles the handovers almost seamlessly and allows the media streaming parameters to be changed to the optimal ones of a new active connection, we can achieve a sub-optimal best-effort single connection scheme.

5. Conclusions and future work

We have presented a novel, transparent, efficient and open vertical handover mechanism which can be easily integrated with media coding and delivery methods that allow for receiving media streams optimized to network and user device capabilities. Ongoing work includes real integration of these components into a single test environment.

Acknowledgment

This research was funded within the framework of a project financed by the Hungarian National Research and Technology Office under contract No. 2/027/2004. The authors would like to thank Csaba Balogh and András Kocsis for their essential work on this project. Also thanks to Péter Patkó, Zoltán B. Kirchner and other colleagues of the industrial partners in the project consortium.

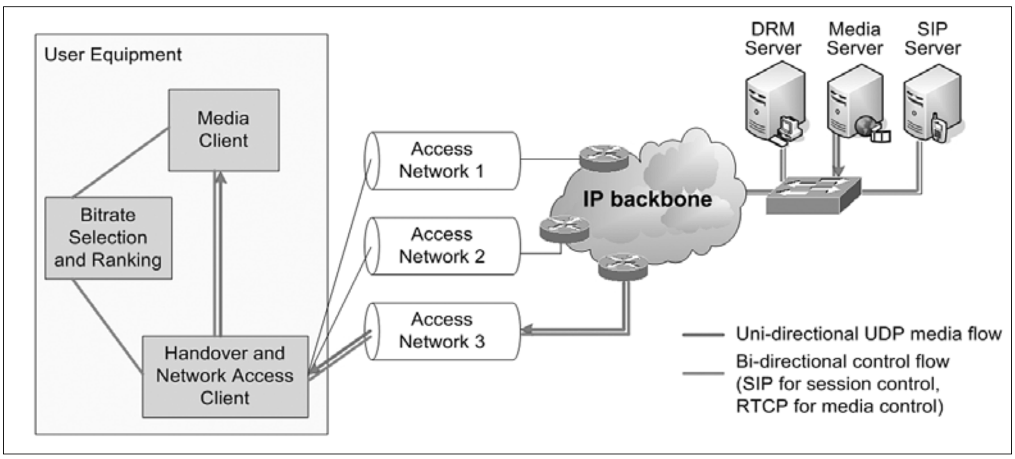


Figure 9. The interconnection of the three main components

References

[1] J. Manner, M. Kojo, “Mobility Related Terminology”, IETF RFC 3753, June 2004.

[2] Cs. A. Szabó, S. Szabó, L. Bokor, “Design Considerations of a Novel Media Streaming Architecture for Heterogeneous Access Environment”, BWAN‘06, September 2006.

[3] S. Doi, S. Ata, H. Kitamura, M. Murata, “Design, Implementation and Evaluation of Routing Protocols for IPv6 Anycast Communication”, AINA‘05, Taiwan, 2005.

[4] Bi-Lynn Ong, Suhaidi Hassan, “A Survey of IPv6 Mobility Management in Real-time Communications”, Networks‘05, v.2, p.6, November 2005.

[5] Q Wang, M. A. Abu-Rgheff, “Integrated Mobile IP and SIP Approach for Advanced Location Management”, 3G ‘03, London, Vol. 494, pp.205–209., June 2003.

[6] J. Rosenberg, H. Schulzrinne, G. Camarillo, R. Sparks, A. Johnston, J. Peterson, M. Handley, E. Schooler, “SIP: Session Initiation Protocol”, IETF RFC 3261, June 2002.

[7] Gonzalo Camarillo, Miguel-Angel Garcia-Martin, “The 3G IP Multimedia Subsystem (IMS): Merging the Internet and the Cellular Worlds”, John Wiley & Sons, 2004.

[8] 3GPP TS 24.229, “Internet Protocol (IP) multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP) Stage 3, (Release 7)”, v7.6.0, December 2006.

[9] Zarri, M., “Future service capabilities offered by the 3GPP system”, 3G Conference 2003, 25-27. June 2003., pp.354–358.

[10] Hector Montes, Gerardo Gomez, Renaud Cuny, Jose F. Paris, “Deployment of IP Multimedia Streaming Services in Third Generation Mobile Networks”, IEEE Wireless Communications, Vol. 9., Nr.5, October 2002., pp.84–92.

[11] DES/TISPAN-02007-NGN-R1, “Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); NGN Functional Architecture Release 1, Overall Architecture”, August 2005.

[12] 3GPP TS 23.228, “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; IP Multimedia Subsystem (IMS) Stage 2 (Release 7)”, v7.6.0, December 2006.

[13] D. Johnson, C. Perkins, J. Arkko, “Mobility Support in IPv6”, IETF RFC 3775, June 2004.